Chapitre 0 : Rappels de probabilités

1 Espace probabilisé

Soit Ω un ensemble. On note $\mathscr{P}(\Omega)$ l'ensemble des parties (=sous-ensembles) de Ω.

Définition 1. Une **tribu** $\mathscr{A}$ sur Ω est un ensemble de parties de Ω qui n'est pas vide, qui est stable par passage au complémentaire et stable par union dénombrable.

Dans toute la suite, on fixe une tribu $\mathscr{A}$ sur Ω.

Définition 2. Une mesure de **probabilité** (ou loi, ou distribution) sur $(\Omega,\mathscr{A})$ est une application $\mathbb{P}:\mathscr{A}\to[0;1]$ telle que $\mathbb{P}(\Omega)=1$ et pour toute suite (A_i) d'ensembles de $\mathscr{A}$, deux à deux disjoints, on a

$$\mathbb{P}\left(\bigcup_{i=0}^{\infty} A_i\right)=\sum_{i=0}^{\infty}\mathbb{P}(A_i).$$

Définition 3. Un **espace probabilisé** est un triplet $(\Omega,\mathscr{A},\mathbb{P})$ où Ω est un ensemble, $\mathscr{A}$ une tribu sur cet ensemble, et $\mathbb{P}$ une mesure de probabilité sur $(\Omega,\mathscr{A})$.

Exemples. Trivial ; ensemble fini ou dénombrable ; mesure uniforme sur $[0;1[$ et la tribu des boréliens ; produits cartésiens.

Dans toute la suite, on fixe un espace probabilisé $(\Omega,\mathscr{A},\mathbb{P})$. Les ensembles de $\mathscr{A}$ seront appelés des **événements**.

Proposition 4. Probabilité d'une union croissante.
Soit $(A_i, i\geq 1)$ une suite d'ensembles de $\mathscr{A}$ telle que pour tout $i\geq 1$, $A_i\subset A_{i+1}$. Alors,

$$\mathbb{P}(\cup_{i=1}^{+\infty}A_i)=\lim_{i\to\infty}\mathbb{P}(A_i).$$

On sait donc calculer la probabilité d'une réunion d'évènements lorsque ces évènements sont disjoints ou emboîtés. Dans les autres cas, on utilise **la formule du crible :**

Proposition 5. *Soit $(A_i, 1\leq i\leq n)$ une suite d'ensembles de $\mathscr{A}$.*

$$\mathbb{P}(\cup_{i=1}^{n}A_i)=\sum_{k=1}^{n}\sum_{1\leq i_1<\cdots<i_k\leq n}(-1)^{k-1}\mathbb{P}(A_{i_1}\cap\cdots\cap A_{i_k}).$$

Dans le cas $n=2$, cette formule devient : $\mathbb{P}(A\cup B)=\mathbb{P}(A)+\mathbb{P}(B)-\mathbb{P}(A\cap B)$.

Définition 6. Soient A et B deux événements tels que $\mathbb{P}(B)>0$. On appelle **probabilité conditionnelle de A sachant B** la quantité

$$\mathbb{P}(A|B)=\frac{\mathbb{P}(A\cap B)}{\mathbb{P}(B)}.$$

L'application $A\in\mathscr{A}\mapsto\mathbb{P}(A|B)$ est une probabilité sur $(\Omega,\mathscr{A})$ telle que $\mathbb{P}(B|B)=1$.

On a donc $\mathbb{P}(A\cap B)=\mathbb{P}(A|B)\mathbb{P}(B)$. Cette formule se généralise au cas de l'intersection de n évènements :

Proposition 7. formule des probabilités composées.
Soit $(A_i, 1\leq i\leq n)$ une suite d'ensembles de $\mathscr{A}$ telle que $\mathbb{P}(A_1\cap\cdots\cap A_{n-1})>0$.

$$\mathbb{P}(A_1\cap\cdots A_n)=\mathbb{P}(A_n|A_1\cap\cdots A_{n-1})\mathbb{P}(A_{n-1}|A_1\cap\cdots A_{n-2})\cdots\mathbb{P}(A_2|A_1)\mathbb{P}(A_1).$$

Proposition 8. formule des probabilités totales.
Soit $(A_i, i\geq 1)$ une partition de Ω (c'est-à-dire une famille d'événements deux à deux disjoints tels que $\cup_i A_i=\Omega$) telle que $\forall i\geq 1$, $\mathbb{P}(A_i)>0$. Pour tout évènement A, on a

$$\mathbb{P}(A)=\sum_{i=1}^{+\infty}\mathbb{P}(A|A_i)\mathbb{P}(A_i).$$

2 Variables aléatoires.

Soit $(\Omega, \mathscr{A})$ un espace muni d'une tribu.

2.1 Définition d'une variable aléatoire.

Définition 9. Cas discret.
Soit E un ensemble fini ou dénombrable, muni de la tribu $\mathscr{E} = \mathscr{P}(E)$. Une variable aléatoire X à valeurs dans E est une fonction $X : \Omega \to E$ telle que, pour tout $x \in E$, $\{X = x\}$ est un événement (i.e. un élément de $\mathscr{A}$).

Notation : $\{X = x\} = \{\omega \in \Omega : \ X(\omega) = x\}$.

Proposition 10. *Soit X une fonction de Ω à valeurs dans E fini ou dénombrable. X est une variable aléatoire ssi pour tout $A \in \mathscr{E}$, $\{X \in A\}$ est un événement.*

Définition 11. Cas réel.
Soit $E = \mathbb{R}$ la droite réelle, munie de $\mathscr{B}$, la tribu des boréliens. Une variable aléatoire réelle est une fonction $X : \Omega \to \mathbb{R}$ telle que, pour tout $A \in \mathscr{B}$, $\{X \in A\}$ est un événement.

Proposition 12. *Soit X une fonction de $(\Omega, \mathscr{A})$ à valeurs dans $(\mathbb{R}, \mathscr{B})$. X est une variable aléatoire ssi pour tout I intervalle ouvert de $\mathbb{R}$, $\{X \in I\}$ est un événement.*

Définition 13. Cas général.
Une variable aléatoire à valeurs dans un ensemble E muni d'une tribu $\mathscr{E}$ est une fonction $X : \Omega \to E$ telle que, pour tout $A \in \mathscr{E}$, $\{X \in A\}$ est un événement.

2.2 Loi d'une variable aléatoire.

Définition 14. Soit $X : (\Omega, \mathscr{A}) \to (E, \mathscr{E})$ une variable aléatoire. L'application $A \in \mathscr{E} \to \mu(A) = \mathbb{P}(X \in A)$ définit une probabilité sur $(E, \mathscr{E})$, appelée **loi de** X.

Notation : $X \sim \mu$ se lit X suit la loi μ.

2.2.1 Cas discret.

Définition 15. Si E est un ensemble fini ou dénombrable, on appelle **fonction de masse** d'une mesure de probabilité μ sur E la fonction $f : E \to [0;1]$ définie, pour tout $x \in E$, par $f(x) = \mu\Big(\{x\}\Big)$. La fonction de masse f_X d'une variable aléatoire X à valeurs dans E est la fonction de masse de sa loi $(f_X(x) = \mathbb{P}(X = x))$.

Proposition 16. *Une mesure de probabilité μ sur un ensemble fini ou dénombrable E est entièrement caractérisée par sa fonction de masse.*

Proposition 17. *Soit E fini ou dénombrable. Une fonction $f : E \to [0;1]$ est la fonction de masse d'une mesure de probabilité sur E si et seulement si, $\sum_{x \in E} f(x) = 1$.*

2.2.2 Cas réel.

Définition 18. On appelle **fonction de répartition** d'une mesure de probabilité μ sur $\mathbb{R}$ la fonction $F : \mathbb{R} \to [0;1]$ définie, pour tout $x \in \mathbb{R}$, par $F(x) = \mu\Big(]-\infty; x]\Big)$. La fonction de répartition F_X d'une variable aléatoire réelle X est la fonction de répartition de sa loi $(F_X(x) = \mathbb{P}(X \leq x))$.

Proposition 19. *Une mesure de probabilité sur $(\mathbb{R}, \mathscr{B})$ est entièrement caractérisée par sa fonction de répartition.*

Théorème 20. *Une fonction $F : \mathbb{R} \to [0;1]$ est une fonction de répartition si et seulement si elle est croissante, continue à droite, $\lim_{x \to -\infty} F(x) = 0$ et $\lim_{x \to +\infty} F(x) = 1$.*

2.3 Densité d'une variable aléatoire.

Définition 21. Une mesure de probabilité μ sur $\mathbb{R}$ admet pour **densité** la fonction $f:\mathbb{R}\to[0;\infty[$ si, pour tout A borélien de $\mathbb{R}$,

$$\mu(A)=\int_{x\in A} f(x)\,\mathrm{d}x.$$

Notations : $X\sim f$ se lit X suit la loi de densité f.

Proposition 22. *Pour qu'une densité existe et soit la fonction f, il suffit de vérifier l'égalité ci-dessus pour tous les ensembles A de la forme $]-\infty;a]$, avec $a\in\mathbb{R}$.*

Proposition 23. *La densité, si elle existe, caractérise entièrement la mesure de probabilité.*

Proposition 24. *Si μ admet une densité et si F est la fonction de répartition de μ, alors $f=F'$ est la densité de μ.*

Proposition 25. *Toute fonction $f:\mathbb{R}\to[0;\infty[$ telle que $\int_{\mathbb{R}} f(x)\,\mathrm{d}x=1$ est une densité d'une certaine mesure de probabilité sur $\mathbb{R}$.*

2.4 Exemples.

Cas discret : lois de Dirac, lois de Bernoulli, binomiales, lois de Poisson, lois géométriques.
Cas réel : lois exponentielles, lois normales.

3 Espérance.

Ici, $(\Omega,\mathscr{A},\mathbb{P})$ désigne un espace probabilisé, et X une variable aléatoire définie sur $(\Omega,\mathscr{A},\mathbb{P})$ à valeurs dans $\mathbb{R}$. L'objectif est de donner un sens mathématique à la définition intuitive : « L'espérance de X est la moyenne des valeurs de X ».

3.1 Fonction indicatrice.

Définition 26. A tout évènement $A\in\mathscr{A}$, on associe la variable aléatoire appelée **indicatrice de** A, et notée $\mathbb{1}_A$, qui est définie par

$$\mathbb{1}_A(\omega)=\begin{cases}1 & \text{si } \omega\in A,\\ 0 & \text{sinon.}\end{cases}$$

La loi de la variable $\mathbb{1}_A$ est donc la loi de Bernoulli de paramètre $\mathbb{P}(A)$. La notation $\mathbb{1}$ remplace donc des conditions (des "si") par des valeurs numériques (0 ou 1). C'est très utile en pratique : on peut additionner et multiplier des nombres, c'est plus difficile de le faire avec des "si". Ainsi,

Proposition 27. *Si A et B sont deux évènements,*
- $\mathbb{1}_{A\cap B}=\mathbb{1}_A\mathbb{1}_B$.
- $\mathbb{1}_{\bar{A}}=1-\mathbb{1}_A$. *Plus généralement, si $(A_i)_{i\in I}$ est une suite d'évènements qui forment une partition de Ω, $\sum_{i\in I}\mathbb{1}_{A_i}=1$.*
- *si A et B sont disjoints, $\mathbb{1}_{A\cup B}=\mathbb{1}_A+\mathbb{1}_B$. Plus généralement, $\mathbb{1}_{A\cup B}=\mathbb{1}_A+\mathbb{1}_B-\mathbb{1}_{A\cap B}$.*

Définition 28. Soit $(A_i)_{i\in I}$ une suite d'évènements qui forment une partition de Ω. On dit que la variable X se décompose sur la partition $(A_i)_{i\in I}$, ssi X peut s'écrire

$$X=\sum_{i\in I}x_i\mathbb{1}_{A_i}, \tag{1}$$

ce qui signifie que X est constante sur chacun des A_i, et vaut x_i sur A_i.

Proposition 29. *Si X prend un nombre fini ou dénombrables de valeurs* **distinctes** $(x_i)_{i\in I}$, *les évènement $A_i=\{X=x_i\}$ forment une partition de Ω sur laquelle X se décompose :*

$$X=\sum_{i\in I}x_i\mathbb{1}_{A_i}.$$

⚠L'écriture (1) n'est pas unique. Par exemple, si A,B,C est une partition de Ω, et $X=2\mathbb{1}_A+\mathbb{1}_B+2\mathbb{1}_C$, on a aussi $X=2\mathbb{1}_{A\cup C}+\mathbb{1}_B$. Cette écriture est unique (à permutation près) si les x_i sont tous distincts.

Définition 30. X est une variable **étagée** ssi il existe une partition $(A_i)_{i\in I}$ (I de cardinal fini ou dénombrable) sur laquelle X se décompose : $X=\sum_{i\in I}\alpha_i\mathbb{1}_{A_i}$. Autrement dit, une variable étagée est une variable dont les valeurs forment un ensemble de cardinal fini ou dénombrable.

3.2 Cas où Ω est de cardinal fini ou dénombrable.

Lorsque l'univers Ω est fini ou dénombrable, la construction de l'espérance est simple à comprendre, et détaillée ci-dessous.

Définition 31. Soit X une variable **positive** définie sur $(\Omega, \mathscr{A}, \mathbb{P})$. L'espérance de X est la quantité

$$\mathbb{E}(X) = \sum_{\omega\in\Omega} X(\omega)\mathbb{P}(\{\omega\}) \in [0;+\infty]. \tag{2}$$

Dans le cas où Ω est un ensemble dénombrable, $\mathbb{E}(X)$ est une série à termes positifs, dont la valeur peut être $+\infty$.

Cas particulier de (2) : pour tout $A \in \mathscr{E}$, $\mathbb{E}(\mathbb{1}_A) = \mathbb{P}(A)$.

Proposition 32. *Soient X et Y deux variables* **positives** *définies sur $(\Omega, \mathscr{A}, \mathbb{P})$.*

1. $\mathbb{E}(X) \geq 0$
2. $\mathbb{E}(X) = 0$ *ssi* $X = 0$ *p.s.*
3. *Si* $0 \leq X \leq Y$ *p.s.*, $0 \leq \mathbb{E}(X) \leq \mathbb{E}(Y)$.
4. *Si α et β sont des réels positifs,* $\mathbb{E}(\alpha X + \beta Y) = \alpha\mathbb{E}(X) + \beta\mathbb{E}(Y)$.

Proposition 33. Ecriture de l'espérance en fonction de la loi de X.
Soit X une variable positive définie sur $(\Omega, \mathscr{A}, \mathbb{P})$, et soit $(x_i)_{i\in I}$ les valeurs distinctes prises par X.

$$\mathbb{E}(X) = \sum_{i\in I} x_i \mathbb{P}(X = x_i) \in [0;+\infty]. \tag{3}$$

Si f est une fonction mesurable de $(\mathbb{R}^+, \mathscr{B}(\mathbb{R}^+))$ dans $(\mathbb{R}^+, \mathscr{B}(\mathbb{R}^+))$,

$$\mathbb{E}(f(X)) = \sum_{i\in I} f(x_i)\mathbb{P}(X = x_i) \in [0;+\infty].$$

Proposition 34. Cas des variables étagées positives.
Soit X est une variable étagée positive : $X = \sum_{i\in I} \alpha_i \mathbb{1}_{A_i}$, où les α_i sont des réels positifs et les $(A_i)_{i\in I}$ forment une partition de Ω,

$$\mathbb{E}(X) = \sum_{i\in I} \alpha_i \mathbb{P}(A_i) \in [0;+\infty]. \tag{4}$$

Soit f est une fonction mesurable de $(\mathbb{R}^+, \mathscr{B}(\mathbb{R}^+))$ dans $(\mathbb{R}^+, \mathscr{B}(\mathbb{R}^+))$,

$$\mathbb{E}(f(X)) = \sum_{i\in I} f(\alpha_i)\mathbb{P}(A_i) \in [0;+\infty].$$

3.3 Cas général.

Sauf dans quelques cas jouets, l'univers Ω est beaucoup plus gros qu'un ensemble de cardinal fini ou dénombrable. Dans ce cadre général, une idée de la construction de l'espérance est donnée en annexe à ce chapitre. Elle suit les étapes suivantes :

- On commence par définir l'espérance d'une variable positive. Cette espérance est un nombre positif pouvant valoir $+\infty$.
- On définit ensuite l'espérance d'une variable X quelconque par $\mathbb{E}(X) = \mathbb{E}(X_+) - \mathbb{E}(X_-)$, où X_+ et X_- sont les parties positives et négatives de X. Cette expression n'a pas de sens si $\mathbb{E}(X_+) = \mathbb{E}(X_-) = +\infty$. Aussi, on ne définit $\mathbb{E}(X)$ que dans le cas où $\mathbb{E}(|X|) < +\infty$. Comme $\mathbb{E}(|X|) = \mathbb{E}(X_+) + \mathbb{E}(X_-)$, cela équivaut à dire que $\mathbb{E}(X_+) < +\infty$ et $\mathbb{E}(X_-) < +\infty$. On dit alors que la variable X est **intégrable.**

Nous nous concentrons ici sur les propriétés importantes de l'espérance.

Proposition 35. Propriétés de l'espérance.
Soient X et Y des variables aléatoires définies sur $(\Omega, \mathscr{A}, \mathbb{P})$ à valeurs dans $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$.

1. *Si* $X \geq 0$ *p.s.*, $\mathbb{E}(X) \geq 0$. *Si* $X \geq 0$ *p.s. et* $\mathbb{E}(X) = 0$, *alors* $X = 0$ *p.s..*
2. *Si X et Y sont positives ou intégrables et $X \leq Y$ p.s.,* $\mathbb{E}(X) \leq \mathbb{E}(Y)$.
3. *Si X et Y sont intégrables, pour tous réels α, β,* $\mathbb{E}(\alpha X + \beta Y) = \alpha\mathbb{E}(X) + \beta\mathbb{E}(Y)$.

Proposition 36. Inégalités.

1. **Inégalité de Jensen.** *Si X est intégrable et φ est une fonction* **convexe** *telle que $\varphi(X)$ est intégrable, alors*

$$\varphi\Big(\mathbb{E}(X)\Big) \leq \mathbb{E}\Big(\varphi(X)\Big).$$

2. **Inégalité de Markov.** *Soit X une variable aléatoire positive. Pour tout $x \in \mathbb{R}^+$, $\mathbb{P}(X \geq x) \leq \frac{1}{x}\mathbb{E}(X)$.*
3. **Inégalité de Bienaymé-Tchebycheff.**
Soit X une variable aléatoire telle que $\mathbb{E}(X^2) < \infty$. On pose $m = \mathbb{E}(X)$, et $\sigma^2 = \mathbb{E}((X-m)^2)$ la moyenne et la variance de X. Pour tout $\epsilon > 0$,
$$\mathbb{P}(|X-m| \geq \epsilon) \leq \frac{\sigma^2}{\epsilon^2}.$$

3.4 Calcul d'espérances.

Proposition 37. Calcul de l'espérance d'une variable discrète.
Soient X une variable aléatoire définie sur $(\Omega, \mathscr{A}, \mathbb{P})$ dont les valeurs $(x_i)_{i\in I}$ forment un ensemble fini ou dénombrable dans $(E, \mathscr{E})$. Soit g une fonction mesurable de $(E, \mathscr{E})$ à valeurs dans $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$ telle que $\sum_{i\in I} |g(x_i)| \mathbb{P}(X = x_i) < +\infty$.
$$\mathbb{E}(g(X)) = \sum_{i\in I} g(x_i)\mathbb{P}(X = x_i).$$

Proposition 38. Calcul de l'espérance d'une variable réelle à densité.
Soit X une variable aléatoire définie sur $(\Omega, \mathscr{A}, \mathbb{P})$ à valeurs $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$, de densité f_X. Soit g une fonction mesurable de $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$ à valeurs dans $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$ telle que $\int_{\mathbb{R}} |g(x)| f_X(x)\, dx < +\infty$.
$$\mathbb{E}(g(X)) = \int_{\mathbb{R}} g(x) f_X(x)\, dx.$$

3.5 Fonction génératrice, fonction caractéristique.

Définition 39. Fonction génératrice.
Soit X une variable aléatoire à valeurs dans $\mathbb{N}$. La **fonction génératrice** de X, notée G_X est la fonction
$$G_X : [-1;1] \to \mathbb{R} \tag{5}$$
$$s \mapsto G_X(s) = \mathbb{E}(s^X) = \sum_{k=0}^{+\infty} s^k \mathbb{P}(X = k) \tag{6}$$

Proposition 40. Propriétés de la fonction génératrice.
1. *La fonction génératrice caractérise la loi de X. $\mathbb{P}(X = k)$ est le coefficient de s^k dans le développement en série entière de G_X.*
$$\mathbb{P}(X = k) = \frac{G_X^{(k)}(0)}{k!}.$$
2. *Si $\mathbb{E}(|X|) < +\infty$, $\mathbb{E}(X) = \lim_{s\to 1^-} G'_X(s)$.*

Définition 41. Fonction caractéristique.
Soit X un vecteur aléatoire à valeurs dans $\mathbb{R}^d$. La **fonction caractéristique** de X, notée ϕ_X est la fonction
$$\phi_X : \mathbb{R}^d \to \mathbb{R} \tag{7}$$
$$t \mapsto \phi_X(t) = \mathbb{E}\left(e^{it.X}\right) = \mathbb{E}\left(e^{i(t_1X_1 + \cdots + t_dX_d)}\right) \tag{8}$$

Proposition 42. Propriétés de la fonction caractéristique.
1. *La fonction caractéristique de X caractérise la loi de X.*
2. *Si $\mathbb{E}(|X|) < +\infty$, ϕ_X est dérivable, et $\phi'_X(0) = i\mathbb{E}(X)$.*

4 Suites de variables aléatoires. Convergences.

Dans tout le chapitre, $(\Omega, \mathscr{A}, \mathbb{P})$ désigne un espace probabilisé. $(X_n)_{n\geq 0}$, X sont des variables aléatoires définies sur $(\Omega, \mathscr{A}, \mathbb{P})$ à valeurs dans $\mathbb{R}^d$.

4.1 Différentes notions de convergence.

Définition 43. Convergence en probabilité.
La suite $(X_n)_{n\geq 0}$ converge en probabilité vers X ($X_n \xrightarrow[n\to+\infty]{proba} X$) ssi pour tout $\epsilon > 0$,

$$\lim_{n\to+\infty} \mathbb{P}\left[\|X_n - X\| > \epsilon\right] = 0.$$

Définition 44. Convergence p.s.
La suite $(X_n)_{n\geq 0}$ converge p.s. vers X ($X_n \xrightarrow[n\to+\infty]{p.s.} X$) ssi

$$\mathbb{P}\left[\left\{\omega; \lim_{n\to+\infty} X_n(\omega) \text{ existe et vaut } X(\omega)\right\}\right] = 1.$$

Définition 45. Convergence dans L^p ($p \geq 1$).
La suite $(X_n)_{n\geq 0}$ converge vers X dans L^p ($X_n \xrightarrow[n\to+\infty]{L^p} X$) ssi

$$\lim_{n\to+\infty} \mathbb{E}\left[\|X_n - X\|^p\right] = 0.$$

Définition 46. Convergence en loi.
La suite $(X_n)_{n\geq 0}$ converge en loi vers X ($X_n \xrightarrow[n\to+\infty]{loi} X$) ssi pour toute fonction f continue bornée de $\mathbb{R}^d$ à valeurs dans $\mathbb{R}$,

$$\lim_{n\to+\infty} \mathbb{E}[f(X_n)] = \mathbb{E}[f(X)].$$

Remarque : La convergence en loi n'est pas à proprement parler une convergence de variables aléatoires. Cette notion de convergence ne concerne que la suite **des lois des** X_n. Ainsi, la définition de la convergence en loi a du sens même si les variables X_n ne sont pas définies sur le même espace $(\Omega, \mathscr{A}, \mathbb{P})$, mais sur des espaces différents $(\Omega_n, \mathscr{A}_n, \mathbb{P}_n)$. En revanche, l'espace des valeurs prises par ces variables est le même pour toutes les variables.

4.2 Attention avec la convergence en loi !

Il faut faire attention à certains "automatismes" quand on travaille avec des convergences en loi. Par exemple, un énoncé du type :

$$\text{"Si } X_n \xrightarrow[n\to+\infty]{\star} X \text{ et } Y_n \xrightarrow[n\to+\infty]{\star} Y, \text{ alors } X_n + Y_n \xrightarrow[n\to+\infty]{\star} X + Y\text{"}$$

est vrai si $\star$ désigne la convergence en probabilité, ou la convergence p.s., ou la convergence dans L^p. **Il n'est plus vrai quand $\star$ désigne la convergence en loi !**

4.3 Limite p.s. et espérance.

La proposition suivante donne des résultats généraux et importants liant limite des espérance et espérance de la limite.

Proposition 47. .

1. **Théorème de convergence monotone, ou de Beppo-Levi.**
   Si (X_n) est une suite croissante de v.a. positives, et si $X = \lim_{n\to+\infty} X_n$ p.s., alors $\lim_{n\to+\infty} \mathbb{E}(X_n) = \mathbb{E}(X)$.
2. **Lemme de Fatou.**
   Si (X_n) est une suite de v.a. positives, alors $\mathbb{E}(\liminf_n X_n) \leq \liminf_n \mathbb{E}(X_n)$.
3. **Théorème de convergence dominée de Lebesgue.**
   Si (X_n) est une suite de v.a. qui converge presque sûrement vers une v.a. X, et s'il existe Y intégrable telle que pour tout n, $|X_n| \leq Y$, alors X est intégrable et $\mathbb{E}\Big(|X_n - X|\Big) \to 0$. En particulier, $\lim_{n\to+\infty} \mathbb{E}(X_n) = \mathbb{E}(X)$.

4.4 Liens entre les différentes notions de convergence.

Proposition 48. *La convergence dans L^p implique la convergence en probabilité.*

Démonstration. Par l'inégalité de Markov, pour tout $\epsilon > 0$, et pour tout $p \geq 1$,

$$\mathbb{P}\left[\|X_n - X\| \geq \epsilon\right] = \mathbb{P}\left[\|X_n - X\|^p \geq \epsilon^p\right] \leq \frac{1}{\epsilon^p}\mathbb{E}\left[\|X_n - X\|^p\right].$$

□

Proposition 49. *La convergence p.s. implique la convergence en probabilité.*

Démonstration. Supposons que $X_n \xrightarrow[n\to+\infty]{p.s.} X$. En notant $Y_n = \mathbb{1}_{\|X_n - X\|>\epsilon}$, $\mathbb{P}\left[\|X_n - X\| > \epsilon\right] = \mathbb{E}(Y_n)$. $Y_n \xrightarrow[n\to+\infty]{p.s.} 0$, et $|Y_n| \leq 1$. Comme 1 est intégrable, on déduit du théorème de convergence dominée que $\mathbb{E}(Y_n) \xrightarrow[n\to+\infty]{} 0$.

□

Proposition 50. *La convergence en probabilité implique la convergence en loi.*

Démonstration. Supposons que $X_n \xrightarrow[n\to+\infty]{proba} X$. Soit f une fonction continue bornée. Pour tout $\epsilon > 0$ $(\epsilon \leq 1)$ et tout $K > 0$,

$$\begin{aligned}
|\mathbb{E}(f(X_n)) - \mathbb{E}(f(X))| &\leq \mathbb{E}\left[|f(X_n) - f(X)|\right] \\
&= \mathbb{E}\left[|f(X_n) - f(X)|\mathbb{1}_{\|X_n - X\|>\epsilon}\right] + \mathbb{E}\left[|f(X_n) - f(X)|\mathbb{1}_{\|X_n - X\|\leq\epsilon}\mathbb{1}_{\|X\|\leq K}\right] + \mathbb{E}\left[|f(X_n) - f(X)|\mathbb{1}_{\|X_n - X\|\leq\epsilon}\mathbb{1}_{\|X\|>K}\right] \\
&\doteq T_1 + T_2 + T_3. \qquad (9)
\end{aligned}$$

$$T_1 \leq 2\|f\|_\infty \mathbb{P}\left[\|X_n - X\| > \epsilon\right].$$

Si $\|X_n - X\| \leq \epsilon$ et $\|X\| \leq K$, on a pour tout n, $\|X_n\| \leq K + \epsilon \leq K + 1$. Ainsi

$$T_2 \leq \sup_{x,y\in C_K;\|x-y\|\leq\epsilon} |f(x) - f(y)|,$$

où $C_K = \left\{x \in \mathbb{R}^d; \|x\| \leq K + 1\right\}$ est compact dans $\mathbb{R}^d$.

Enfin,

$$T_3 \leq 2\|f\|_\infty \mathbb{P}\left[\|X\| > K\right].$$

On passe à la limite quand $n \to +\infty$ dans (9). Comme $X_n \xrightarrow[n\to+\infty]{proba} X$, T_1 tend vers 0, et on obtient que pour tout $\epsilon > 0$ $(\epsilon \leq 1)$ et tout $K > 0$,

$$\limsup_{n\to+\infty} |\mathbb{E}(f(X_n)) - \mathbb{E}(f(X))| \leq \sup_{x,y\in C_K;\|x-y\|\leq\epsilon} |f(x) - f(y)| + 2\|f\|_\infty \mathbb{P}\left[\|X\| > K\right].$$

On passe maintenant à la limite quand $\epsilon \to 0$. La fonction f étant continue, elle est uniformément continue sur le compact C_K, et $\sup_{x,y\in C_K;\|x-y\|\leq\epsilon} |f(x) - f(y)|$ tend vers 0 quand $\epsilon \to 0$. On obtient donc que pour tout $K > 0$,

$$\limsup_{n\to+\infty} |\mathbb{E}(f(X_n)) - \mathbb{E}(f(X))| \leq 2\|f\|_\infty \mathbb{P}\left[\|X\| > K\right].$$

Enfin, on passe à la limite quand $K \to +\infty$ pour conclure. □

En résumé,

$$\begin{array}{ccccc}
\text{p.s.} & & & & \\
 & \searrow & & & \\
 & & \text{probabilité} & \Rightarrow & \text{loi} \\
 & \nearrow & & & \\
L^p & & & &
\end{array}$$

4.5 Deux critères de convergence en loi.

Proposition 51. Convergence en loi et fonctions de répartition.
Soit $(X_n)_{n\geq 1}$ et X des variables aléatoires à valeurs dans $\mathbb{R}$ de fonctions de répartition $(F_n)_{n\geq 1}$ et F. Il y a équivalence entre

1. $X_n \xrightarrow[n\to+\infty]{loi} X$.
2. *Pour tout réel t où F est continue,* $F_n(t) \xrightarrow[n\to+\infty]{} F(t)$.

Corollaire 52. Soit $(X_n)_{n\geq 1}$ une suite de variables aléatoires à valeurs dans $\mathbb{R}$, et a un réel. Si $X_n \xrightarrow[n\to+\infty]{loi} a$, alors $X_n \xrightarrow[n\to+\infty]{proba} a$.

Pour cette raison, on pourra reprendre nos automatismes quand on "compose" deux convergences en loi, **dès que l'une des deux convergences a lieu vers une constante**. Par exemple, l'énoncé suivant est vrai :

"Soit a un réel. Si $X_n \xrightarrow[n\to+\infty]{loi} a$ et $Y_n \xrightarrow[n\to+\infty]{loi} Y$, alors $X_n + Y_n \xrightarrow[n\to+\infty]{loi} a + Y$."

Proposition 53. Convergence en loi et fonctions caractéristiques.
Soit $(X_n)_{n\geq 1}$ et X des vecteurs aléatoires à valeurs dans $\mathbb{R}^d$ de fonctions caractéristiques $(\phi_n)_{n\geq 1}$ et ϕ. Il y a équivalence entre

1. $X_n \xrightarrow[n\to+\infty]{loi} X$.
2. *Pour tout vecteur $t \in \mathbb{R}^d$,* $\phi_n(t) \xrightarrow[n\to+\infty]{} \phi(t)$.

Remarque : L'implication 1. ⇒ 2. est une conséquence directe de la définition de la convergence en loi.
C'est l'implication 2. ⇒ 1. qui est utile. Cette dernière est une version affaiblie du théorème suivant dû à Paul Lévy :

Théorème 54. Théorème de Paul Lévy.
Soit $(X_n)_{n\geq 1}$ des vecteurs aléatoires à valeurs dans $\mathbb{R}^d$ de fonctions caractéristiques $(\phi_n)_{n\geq 1}$. Si pour tout vecteur $t \in \mathbb{R}^d$, $\phi_n(t)$ converge quand $n \to +\infty$ vers une fonction $\phi(t)$ qui est continue en 0, alors la fonction ϕ est la fonction caractéristique d'un vecteur aléatoire X, et on a $X_n \xrightarrow[n\to+\infty]{loi} X$.

5 Exercices

Exercice 1. Soit $(A_i)_{i\geq 0}$ une suite d'événements telle que $\mathbb{P}(A_i) = 1$ pour tout $i \geq 0$. Montrer que $\mathbb{P}\Big(\bigcap_i A_i\Big) = 1$.

Exercice 2. Démontrer la formule du crible, la formule des probabilités composés, la formule des probabilités totales.

Exercice 3. Calculer la fonction de répartition de (a) la loi binomiale $\mathscr{B}(n,p)$, (b) la loi de Poisson $\mathscr{P}(\lambda)$, (c) la loi exponentielle $\mathscr{E}\mathrm{xp}(\lambda)$, (d) la loi dont la densité est $f(x) = x\exp(-x^2)/2$ si $x \geq 0$ (et zéro sinon).

Exercice 4. Soit une loi de fonction de répartition F. Un nombre m est une médiane de F si $\lim_{y\to m-} F(y) \leq 1/2 \leq F(m)$. Un tel nombre existe-t-il toujours ? Est-il unique ?

Exercice 5. Soit X une variable aléatoire réelle de fonction de répartition F. Pour tout $a,b \in \mathbb{R}$, $a \leq b$, exprimer les quantités $\mathbb{P}(X \in]a,b])$, $\mathbb{P}(X \in]a,b[)$, $\mathbb{P}(X \in [a,b])$, $\mathbb{P}(X \in [a,b[)$, $\mathbb{P}(X = a)$, en fonction de F, a et b.

Exercice 6. Soit X une variable aléatoire réelle de fonction de répartition continue F. Soit G une fonction continue et strictement croissante sur $\mathbb{R}$. Trouver les fonctions de répartition des variables aléatoires ci dessous :
$-X$, X^2, $|X|$, $\sin(X)$, $X^+ = \max(0,X)$, $X^- = -\min(0,X)$, $G^{-1}(X)$, $F(X)$, $G^{-1}\Big(F(X)\Big)$.

Exercice 7. Soit X une variable aléatoire réelle de fonction de répartition F, et $a < b$ deux réels. Calculer les fonctions de répartition des variables aléatoires Y et Z définies par

$$Y = \begin{cases} a & \text{si } X < a, \\ X & \text{si } a \leq X \leq b, \\ b & \text{si } X > b \end{cases} \qquad Z = \begin{cases} X & \text{si } |X| \leq b, \\ 0 & \text{sinon.} \end{cases}$$

Exercice 8. Trouver (si elle existe) une valeur de c telle que la fonction f définie par la formule ci-dessous est une densité.

$$\text{(a) } f(x) = \begin{cases} cx^{-d}, & \text{si } x > 1, \\ 0 & \text{sinon.} \end{cases} \qquad \text{(b) } f(x) = ce^x\Big(1+e^x\Big)^{-2}.$$

Exercice 9. Soit $a>0$ et $0<p<1$. Soit X une variable aléatoire de loi de Poisson $\mathscr{P}(a)$. Soit Y une variable aléatoire à valeurs entières telle que, pour tous entiers $0\le k\le n$, $\mathbb{P}(Y=k|X=n)=\binom{n}{k}p^k(1-p)^{n-k}$. Quelle est la loi de Y ? Et de $Z=X-Y$?

Exercice 10. Soit T une variable à valeurs dans $\mathbb{N}$ telle que
— $\forall n\in\mathbb{N}$, $\mathbb{P}(T\ge n)>0$.
— $\forall n,m\in\mathbb{N}$, $\mathbb{P}(T\ge n+m|T\ge n)=\mathbb{P}(T\ge m)$.

1. T est dite sans mémoire. Pourquoi ? Est-il raisonnable de modéliser la durée d'une conversation téléphonique par une variable sans mémoire ?
2. Montrer que T est une variable géométrique de paramètre $q=\mathbb{P}(T=0)$ (i.e $\forall n\in\mathbb{N}$, $\mathbb{P}(T=n)=(1-q)^nq$).

Exercice 11. 1. Soit T une variable aléatoire sur $\mathbb{R}^+$, telle que
— $\forall t>0$, $\mathbb{P}(T>t)>0$;
— $\forall t,s>0$, $\mathbb{P}(T>t+s|T>s)=\mathbb{P}(T>t)$.
T est dite sans mémoire. Pourquoi ?

2. On définit l'application $f(t)=\log(\mathbb{P}(T>t)),t>0$. Montrer que $f(x)=x.f(1)$ pour tout x rationnel positif, puis pour tout x réel positif. Conclure que T est une variable exponentielle.
3. Soit T une variable exponentielle. Quelle est la loi de la variable $[T]$ (partie entière de T) ?

Exercice 12. Soit X une variable prenant les valeurs 3, -1, 1. Soit Y une variable prenant les valeurs -1, 3. Ecrire les variables X, Y, $\exp(X)$, X^2, $X+Y$ sous la forme $\sum_{i\in I}\alpha_i\mathbb{1}_{A_i}$ où les α_i sont tous distincts.

Exercice 13. Dans un lancer de dé, on gagne deux euros si le nombre tiré est premier, et on en perd deux si le nombre tiré est pair. Ainsi, si on tire le numéro 2, on gagne 2-2=0 euros. On note X le gain obtenu.

1. Ecrire un triplet $(\Omega,\mathscr{A},P)$ modélisant l'expérience.
2. Ecrire X sous la forme $\sum_{i\in I}\alpha_i\mathbb{1}_{A_i}$ où les α_i sont tous distincts.
3. Calculer $\mathbb{E}(X)$.

Exercice 14. Calculer l'espérance d'une variable X : *(1)* de loi binomiale $\mathscr{B}(n,p)$ (Réponse : np) ; *(2)* de loi de Poisson $\mathscr{P}(\lambda)$ (Réponse : λ) ; *(3)* de loi exponentielle $\mathscr{E}(\lambda)$ (Réponse : $1/\lambda$) ; *(4)* de loi normale $\mathscr{N}(\mu,\sigma^2)$ (Réponse : μ).

Exercice 15. Soit X une variable aléatoire à valeurs dans $\mathbb{N}$. On pose pour tout entier j : $p_j=\mathbb{P}(X=j)$ et $q_j=\mathbb{P}(X>j)$. Montrer que l'espérance de X est

$$\mathbb{E}(X)=\sum_{j=0}^{\infty}q_j$$

Exercice 16. Pour $a>0$, on pose $\Gamma(a)=\int_0^{+\infty}e^{-x}x^{a-1}dx$. On appelle loi Gamma de paramètres a et λ ($a>0$ et $\lambda>0$) notée $G(a,\lambda)$ la loi sur $\mathbb{R}$ de densité $\gamma_{a,\lambda}$ où

$$\gamma_{a,\lambda}(x)=\frac{\lambda^a}{\Gamma(a)}e^{-\lambda x}x^{a-1}1_{\mathbb{R}^+}(x).$$

1. Vérifier que $\Gamma(a)$ est défini pour $a>0$, montrer que $\Gamma(a+1)=a\Gamma(a)$ et calculer $\Gamma(n)$ pour $n\in\mathbb{N}^\star$.
2. Soit X une variable aléatoire de loi $G(a,\lambda)$. Calculer $\mathbb{E}(X)$ et $\mathrm{var}(X)$.
3. Soit Y une variable aléatoire de loi gaussienne centrée réduite $\mathscr{N}(0,1)$. Montrer que Y^2 a la loi gamma $G(1/2,1/2)$. En déduire la valeur de $\Gamma(1/2)$.

Exercice 17. Pour n,m entiers naturels, on note $B(n,m)=\int_0^1 x^n(1-x)^m\,dx$. On donne $B(n,m)=\frac{n!m!}{(n+m+1)!}$, et on considère

$$f_{n,m}(x)=\frac{1}{B(n,m)}x^n(1-x)^m\mathbb{1}_{[0,1]}(x).$$

1. Montrer que $f_{n,m}$ est une densité de probabilité.
2. Montrer que si X est de densité $f_{n,m}$, alors pour tout entier p, $\mathbb{E}(X^p)=\frac{(n+m+1)!(n+p)!}{n!(n+m+p+1)!}$.
3. Si X est de densité $f_{n,m}$, quelle est la loi de $1-X$, de $\frac{X}{1-X}$?

Exercice 18. On dit que Z suit la loi Log-normale de paramètres m et σ ($\sigma>0$) si Z est une variable aléatoire à valeurs dans $\mathbb{R}_+^\star$ telle que $\ln(Z)$ est de loi $\mathscr{N}(m,\sigma^2)$.

1. Quelle est la densité de Z ?
2. Déterminez l'espérance et la variance de Z.

Exercice 19. 1. Calculer la fonction génératrice d'une variable X de loi binomiale $\mathscr{B}(n,p)$ (Réponse : $(1-p+sp)^n$).

2. Calculer la fonction génératrice d'une variable X de loi de Poisson $\mathscr{P}(\lambda)$ (Réponse : $\exp(-\lambda(1-s))$).

Exercice 20. 1. Calculer la fonction caractéristique d'une variable X de loi exponentielle $\mathscr{E}(\lambda)$ (Réponse : $\lambda/(\lambda-it)$) ;

2. Calculer la fonction caractéristique d'une variable X de loi exponentielle symétrique de paramètre λ de densité $\frac{\lambda}{2}\exp(-\lambda|x|)$ sur $\mathbb{R}$ (Réponse : $\lambda^2/(\lambda^2+t^2)$).
3. En vous aidant de la question précédente, et en pensant à l'inversion de Fourier, calculer la fonction caractéristique d'une variable X de loi de Cauchy de densité $1/(\pi(1+x^2))$ sur $\mathbb{R}$ (Réponse : $\exp(-|t|)$).
4. Calculer la fonction caractéristique d'une variable X de loi normale $\mathscr{N}(\mu,\sigma^2)$ (Réponse : $\exp(it\mu)\exp\left(-\sigma^2t^2/2\right)$).

Exercice 21. Soit X une variable aléatoire à valeurs dans $\mathbb{N}$ dont la fonction génératrice est définie sur $[-1,1]$ par $G_X(t)=\frac{t^2}{2-t^2}$. Déterminer les lois de X et $Y \coloneqq \frac{X}{2}$.

Exercice 22. Soit $(X_n)_{n\geq 0}$ une suite de variables aléatoires rélles telles que pour tout n, X_n est de densité $f_n(x)=\frac{n}{\pi(1+n^2x^2)}$. Montrer que $X_n \xrightarrow[n\to+\infty]{loi} 0$.

Exercice 23. Convergence de la loi binomiale vers la loi de Poisson.
Soit $(X_n)_{n\geq 0}$ une suite de v.a. avec $X_n \sim \mathscr{B}(n,p_n)$ et $np_n \xrightarrow[n\to+\infty]{} \lambda>0$. Montrez que $X_n \xrightarrow[n\to+\infty]{loi} \mathscr{P}(\lambda)$.

Annexe sur la construction de l'espérance.

Cas des variables étagées positives.
On ne suppose plus que Ω est de cardinal fini ou dénombrable. Dans ce cas, l'expression (2) n'a pas de sens (à cause de la somme sur un nombre infini non dénombrable de termes). En revanche, les expressions (3) et (4) continuent à avoir du sens pour toute variable positive prenant un nombre fini ou dénombrable de valeurs. Ce sont ces expressions qui servent alors de définition de l'espérance.

Si on on choisit l'expression (4) comme définition de l'espérance d'une variable étagée, il faut vérifier que cette expression ne dépend pas de l'écriture choisie (Par exemple, $X=2\mathbb{1}_A+\mathbb{1}_B+2\mathbb{1}_C=2\mathbb{1}_{A\cup C}+\mathbb{1}_B$ amène à deux expressions pour $\mathbb{E}(X)$: $\mathbb{E}(X)=2\mathbb{P}(A)+\mathbb{P}(B)+2\mathbb{P}(C)$ et $\mathbb{E}(X)=2\mathbb{P}(A\cup C)+\mathbb{P}(B)$, qui sont bien identiques). L'expression (3) est alors un cas particulier de (4).

Si on choisit l'expression (3) comme définition, il faut alors montrer que l'expression (4) de l'espérance d'une variable étagée est vraie. Cela peut se faire en utilisant la linéarité de l'espérance, à condition d'avoir démontré cette linéarité à partir de la définition (3)... Essayez de vous convaincre que cette linéarité est bien vérifiée, en regardant par exemple un cas particulier.

Avec cette définition, les propositions 32, 33 et 34 restent vraies.

Cas d'une variable positive quelconque.
Dans le cas où l'ensemble des valeurs prises par X n'est pas dénombrable, aucune des expressions (3) ou (4) n'a de sens. Pour définir $\mathbb{E}(X)$, on utilise un passage à la limite. On approche X par une suite décroissante $(X_n)_{n\geq 1}$ de variables étagées (par exemple, $X_n=\sum_{k=0}^{+\infty}\frac{k+1}{2^n}\mathbb{1}_{\frac{k}{2^n}\leq X<\frac{k+1}{2^n}}$). Comme $0\leq X_{n+1}\leq X_n$, on a $\mathbb{E}(X_{n+1})\leq\mathbb{E}(X_n)$. La suite $(\mathbb{E}(X_n))_{n\geq 1}$ est une suite décroissante de réels positifs. Par conséquent, cette suite admet une limite qu'on définit comme $\mathbb{E}(X)$:

$$\mathbb{E}(X)=\lim_{n\to+\infty}\mathbb{E}(X_n).$$

Pour vérifier qu'il s'agit bien d'une définition, il faut s'assurer que cette limite ne dépend pas de la suite décroissante de variables étagées qui approchent X.

Proposition 55. *La proposition 32 reste vraie.*

Espérance d'une variable réelle quelconque.
Soit X une variable réelle qui n'est pas nécessairement positive. Pour tout réel x, on définit la partie positive x_+ de x par $x_+=\max(x,0)$, et sa partie négative x_- par $x_-=\max(-x,0)$. Ce sont deux nombres positifs, et on a $x=x_+-x_-$ et $|x|=x_++x_-$. On peut donc écrire

$$X=X_+-X_- \qquad \text{et} \qquad |X|=X_++X_-.$$

On a vu comment définir $\mathbb{E}(X_+)$ et $\mathbb{E}(X_-)$. On définit alors $\mathbb{E}(X)$ en posant $\mathbb{E}(X)=\mathbb{E}(X_+)-\mathbb{E}(X_-)$ dès que cette expression a du sens, i.e. dès qu'on ne se retrouve pas avec la forme indéterminée $+\infty-\infty$. C'est en particulier le cas lorsque $\mathbb{E}(X_+)<+\infty$ et $\mathbb{E}(X_-)<+\infty$, ou de façon équivalente lorsque $\mathbb{E}(|X|)=\mathbb{E}(X_+)+\mathbb{E}(X_-)<+\infty$